GRIDS WITH DENSE VALUES

URI SHAPIRA

ABSTRACT. Given a continuous function from Euclidean space to the real line, we analyze (under some natural assumption on the function), the set of values it takes on translates of lattices. Our results are of the flavor: For almost any translate the set of values is dense in the set of possible values. The results are then applied to a variety of concrete examples obtaining new information in classical discussions in different areas in mathematics; in particular, Minkowski's conjecture regarding products of inhomogeneous forms and inhomogeneous Diophantine approximations.

1. INTRODUCTION

Given a continuous function $F : \mathbb{R}^d \to \mathbb{R}$, it is a natural question in number theory to try and analyze the set of values F takes on points of a lattice in $\mathbb{R}^d$. Up to a linear change of variable this question is of course equivalent to analyzing the values F takes at integer points. We shall be interested in an inhomogeneous variant of this discussion; we try to analyze the set of values F takes on grids, that is on translated lattices. We approach this discussion from a dynamical point of view which leads us to impose some natural assumptions on the function F under consideration. We present a variety of concrete examples in §2 and applications to classical discussions in Diophantine approximations and the geometry of numbers in §3. The main tools we develop to derive our results are the Mixing and the Coset lemmas appearing in §4. The discussion in §4 is concerned with closures of certain random sequences on the d-torus. It is independent of the rest of this paper and may be of independent interest on its own.

1.1. **Basic notions.** The basic objects we shall work with are lattices and grids. As we wish to exploit dynamical methods it is most convenient to present spaces the points of which are lattices and grids. We set $G = \mathrm{SL}_d(\mathbb{R}), \Gamma = \mathrm{SL}_d(\mathbb{Z})$, and let $X_d = G/\Gamma$. A point $g\Gamma \in X_d$ will be denoted by $\bar{g}$. The space X_d can be identified with the space of unimodular (i.e. of covolume 1) lattices in $\mathbb{R}^d$. The identification is defined by associating to $\bar{g} \in X_d$ the lattice spanned by the columns of g. We let $Y_d = \left(G \ltimes \mathbb{R}^d \right) / \left(\Gamma \ltimes \mathbb{Z}^d \right)$. The space Y_d is identified with the space of unimodular *grids* in $\mathbb{R}^d$, where by a grid we mean a set of the form $\bar{g}+v = \{w + v : w \in \bar{g}\}$. Points of Y_d are denoted by $\bar{g}+v$ where $g \in G$ and $v \in \mathbb{R}^d$. We endow X_d, Y_d with the quotient topology induced from the usual topology on G and $G \ltimes \mathbb{R}^d$ respectively. In fact, X_d, Y_d are smooth manifolds and inherit a complete Riemannian metric from the corresponding group. It will be convenient to interchangeably think of the points of X_d, Y_d as subsets of $\mathbb{R}^d$ or as points in the corresponding manifold.

We shall denote points of X_d, Y_d by lower case letters x, x_1, y, y' etc. It is a good exercise for the reader who is not familiar with the topologies introduced above, to work out the meaning of two lattices (resp. grids) being close to each other when thinking of them as subsets in $\mathbb{R}^d$. This boils down to saying that in a very large box, centered at the origin, the corresponding two sets are close in the usual sense.

There is an obvious projection $\pi : Y_d \to X_d$ given by $\pi(\bar{g} + v) = \bar{g}$. The fiber $\pi^{-1}(x)$ for $x \in X_d$ is naturally identified with the torus $\mathbb{R}^d/x$. Note also that X_d naturally embeds in Y_d. For a continuous function $F : \mathbb{R}^d \to \mathbb{R}$ we define for each $y \in Y_d$ the *value set*

$$V_F(y) = \{F(v) : v \in y\}. \tag{1.1}$$

In general, it is of interest to analyze for a fixed grid $y \in Y_d$ the value set $V_F(y)$. In particular, one would like to answer questions such as: Is the value set dense, discrete, or does it contain zero in its closure. We make the following definition which will be of most interest to us in this paper.

Definition 1.1. Let $F : \mathbb{R}^d \to \mathbb{R}$ be a function.
  (1) A grid $y \in Y_d$ is DV$_F$ (dense values) if the value set $V_F(y)$ is dense in the image $F(\mathbb{R}^d)$.
  (2) A lattice $x \in X_d$ is grid-DV$_F$ if all the grids $x + v$, $v \in \mathbb{R}^d$ are DV$_F$.
  (3) Given a lattice $x \in X_d$ and a probability measure μ on the torus $\pi^{-1}(x)$, we say that x is μ-almost surely grid-DV$_F$ if μ-almost any grid $y \in \pi^{-1}(x)$ is DV$_F$ (we sometimes express this by saying that x is almost surely grid-DV$_F$ with respect to μ) .

In general, it is a hard problem to decide for a specific function F whether or not there exist grid-DV$_F$ lattices (although in some concrete examples the answer is known, see §2). On the other hand, the existence of almost surely grid-DV$_F$ lattices (with respect to some natural family of measures) is guaranteed once some reasonable assumptions on F are imposed. In this paper we find sufficient conditions for a lattice $x \in X_d$ to be almost surely grid-DV$_F$ with respect to some measures which we now turn to describe.

1.2. **Haar measures of subtori.** Let $x \in X_d$ be given. We say that a subspace $U < \mathbb{R}^d$ is rational with respect to x, if $x \cap U$ is a lattice in U. The rational subspaces are in one to one correspondence with the closed connected subgroups of the torus $\pi^{-1}(x)$. We refer to these as *subtori* and denote the subtorus corresponding to a rational subspace U by $U + x$. The Haar probability measure on such a subtorus is denoted by λ_U. For $v \in \mathbb{R}^d$ and a subtorus $U + x$ we denote by $v + (U + x)$ the coset of v with respect to the subtorus. The translation of the Haar measure λ_U by v, supported on the coset, is denoted by λ_{v+U}. When $U = \mathbb{R}^d$ we sometimes refer to λ_U as Lebesgue measure.

1.3. **The invariance group.** Let

$$\mathbf{H}_F = \{g \in G : F \circ g = F\},$$
$$H_F = \text{ The connected component of the identity in } \mathbf{H}_F.$$

The group H_F will be referred to as the *invariance group* of F. It is a closed subgroup of G and in most interesting cases it is the connected component of the identity of an algebraic group. This happens when F is a polynomial but also in some other cases too. We say that F is *noncompact* if H_F is noncompact . We shall only discuss the noncompact case as we wish to exploit dynamical properties of the H_F-action on X_d, Y_d which are not present in the compact case.

The group G acts naturally by left translation on X_d, Y_d. These actions commute with π. When thinking of the points of X_d, Y_d as subsets of $\mathbb{R}^d$ these actions are induced by the linear action of G on $\mathbb{R}^d$. The importance of the invariance group to our discussion is that it leaves the value sets invariant; for any $y \in Y_d$ and any $h \in H_F$,

$$V_F(y) = V_F(hy). \tag{1.2}$$

The following lemma illustrates how dynamics comes into the game. The reader should deduce it from equation (1.2), the continuity of F, and the topology of Y_d:

Lemma 1.2 (Inheritance). *Let $y, y_0 \in Y_d$ be such that $y_0 \in \overline{H_F y}$. Then $\overline{V_F(y_0)} \subset \overline{V_F(y)}$. In particular, if y_0 is DV_F then so is y. If y has a discrete value set then so does y_0.*

As $SL_d(\mathbb{Z})$ is a lattice in $SL_d(\mathbb{R})$, the space X_d carries a G-invariant probability measure. Similarly, Y_d carries a $G \ltimes \mathbb{R}^d$-invariant probability measure. We shall refer to both these measures as the Haar measures. It is not hard to show that the Haar measure on Y_d disintegrates to the Lebesgue measures on the fibers $\{\pi^{-1}(x) : x \in X_d\}$ with respect to the Haar measure on X_d.

Moore's theorem [19] (see also Theorem 2.2.19 in [24]) asserts that the ergodicity of the G-action on Y_d with respect to the Haar measure, implies the ergodicity of the action of any noncompact subgroup of G. As we assume that H_F is noncompact it acts ergodically on Y_d. It follows in particular, that for almost any grid $y \in Y_d$, $H_F y$ is dense in Y_d. This gives us the following immediate corollary of the inheritance lemma

Corollary 1.3. *Almost any $y \in Y_d$ is DV_F. In other words, almost any lattice $x \in X_d$ is almost surely grid-DV_F with respect to the Lebesgue measure of the full torus $\pi^{-1}(x)$.*

In this paper we wish to point out a fairly general connection between the dynamical behavior of a lattice $x \in X_d$ under the action of the invariance group, and the value set $V_F(y)$ along a fiber $\pi^{-1}(x)$. Our aim is to sharpen Corollary 1.3 and to develop better understanding of the set of almost surely grid-DV_F lattices.

Before ending this introduction let us make the following definitions which we need in order to state our results. We say that a sequence in a topological space is *divergent* if it has no converging subsequences. An orbit $H_F x$ is said to be *divergent* if for any divergent sequence $h_n \in H_F$ the sequence $h_n x$ is divergent in X_d (i.e. if the orbit map $h \mapsto hx$ is proper).

Definition 1.4. A continuous function $F : \mathbb{R}^d \to \mathbb{R}$ is *nondegenerate* if for any subspace $\{0\} \neq U < \mathbb{R}^d$ and any grid $y \in Y_d$, one has

$$\{F(u + v) : u \in U, v \in y\} = F(\mathbb{R}^d). \tag{1.3}$$

A typical example of a degenerate function, which we wish to avoid, is a polynomial with coefficients in $\mathbb{Z}$ which does not depend on one of the variables (the simplest example is $F : \mathbb{R}^2 \to \mathbb{R}$ given by $F(v_1, v_2) = v_1$). In this case, choosing the grid to be simply $\mathbb{Z}^d$ and U to be the line corresponding to the variable that does not appear in F, we see that the set in (1.3) is discrete.

Definition 1.5. A sequence $h_n \in G$ is said to be *almost finite* with respect to a subspace $U < \mathbb{R}^d$ if there exist a sequence $\epsilon_n \in G$ with $\epsilon_n \to e$ (e being the identity element in G), such that the set of restrictions $\{\epsilon_n h_n|_U\}$ is finite.

Note that any diverging sequence h_n is not almost finite with respect to $\mathbb{R}^d$. Also, if h_n are diagonalizable and the eigenvalues approach either 0 or ∞, then h_n is not almost finite with respect to any nontrivial subspace.

2. EXAMPLES AND RESULTS

2.1. **Main theorem.** The following theorem is a simplified version of our main result. This version is sufficiently strong for most of our applications.

Theorem 2.1. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be nondegenerate and noncompact. Let $x \in X_d$ be a lattice with a nondivergent H_F-orbit and let λ be the Lebesgue measure of the full torus $\pi^{-1}(x)$. Then x is λ-almost surely grid-DV_F.*

Theorem 2.1 follows from the following theorem which is the main result in this paper. It is proved in §5.

Theorem 2.2. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be nondegenerate and noncompact. Let $x \in X_d$ be a lattice with a nondivergent H_F-orbit. Let $U < \mathbb{R}^d$ be a subspace, rational with respect to x, and $w \in \mathbb{R}^d$. Finally, let λ_{w+U} be the translation of the Haar measure supported on the subtorus $U + x$ by w. Then, if there exists a divergent sequence $h_n \in H_F$ such that $h_n x$ converges and h_n is not almost finite with respect to U, then x is λ_{w+U}-almost surely grid-DV_F.*

In some examples (as will be seen below) the fact that the sequence h_n is not almost finite follows automatically from properties of the invariance group. We now turn to apply these results to a variety of concrete examples.

2.2. **Rank one.** The following family of examples is particularly relevant to Diophantine approximations (see §3.2). Let m, n be positive integers such that $m + n = d$. Let us denote vectors in $\mathbb{R}^d$ as column vectors $(\mathbf{x}, \mathbf{y})^t$ (here t stands for transpose), where $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m$. Let $P_{n,m} : \mathbb{R}^d \to \mathbb{R}$ be defined by $P_{n,m}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|_2^n \|\mathbf{y}\|_2^m$ (where we denote here by $\|\cdot\|_2$ the Euclidean norm on the corresponding space). Denoting the invariance group of $P_{n,m}$ by $H_{n,m}$, the reader should verify that $H_{n,m} = \mathrm{SO}(n) \times \mathrm{SO}(m) \times \{a_{n,m}(t)\}_{t \in \mathbb{R}}$, where

$$a_{n,m}(t) = \mathrm{diag}\left(\underbrace{e^{mt}, \ldots, e^{mt}}_{n}, \underbrace{e^{-nt}, \ldots, e^{-nt}}_{m}\right). \tag{2.1}$$

The reader should verify that $P_{n,m}$ is nondegenerate and noncompact. Hence, Theorem 2.2 applies. In fact, we have the following theorem.

Theorem 2.3. *Let $x \in X_d$ be a lattice with a nondivergent $H_{n,m}$-orbit. Then for any subspace $\{0\} \neq U < \mathbb{R}^d$, rational with respect to x, and any $w \in \mathbb{R}^d$, x is λ_{w+U}-almost surely grid-$\mathrm{DV}_{P_{n,m}}$.*

Proof. Let λ_{w+U} be a measure as in the statement. As the orbit $H_{n,m}x$ is nondivergent, there exists a diverging sequence $h_n \in H_F$ such that $h_n x$ converges. In fact, as $H_{n,m}$ is the product of a compact group with the one parameter group $a_{n,m}(t)$, we may assume that $h_n = a_{n,m}(t_n)$ for some $t_n \to \pm\infty$. As the eigenvalues of $a_{n,m}(t_n)$ approach 0 or ∞, it follows that h_n is not almost finite with respect to U. Theorem 2.2 applies and we conclude that x is λ_{w+U}-almost surely grid-$\mathrm{DV}_{P_{n,m}}$. $\qquad\square$

In §3.2 we apply Theorem 2.3 to Diophantine approximations and in particular, to derive new results on nonsingular forms. We also remark here that for any lattice $x \in X_d$ there always exist grids which are not $\mathrm{DV}_{P_{n,m}}$; that is, there are no grid-$\mathrm{DV}_{P_{n,m}}$ lattices in this case. This result was proved by Davenport [9] for the case $d = 2, n, m = 1$ and by Einsiedler and Tseng for general d, $n = d - 1, m = 1$ [10]. In fact, in [10] it is shown that the set of grids which are not $\mathrm{DV}_{P_{n,m}}$ is of full Hausdorff dimension (and moreover, a winning set for Schmidt's game). See also [5] for generalizations. The reason behind this phenomenon seems to be the fact that the invariance group does not contain any unipotents and is of real rank 1.

2.3. **The product of linear forms.** Let $N : \mathbb{R}^d \to \mathbb{R}$ be the function $N(x) = \prod_1^d x_i$. We denote the invariance group H_N by A. It is the group of diagonal matrices with positive diagonal entries in G. The reader should verify that N is nondegenerate and noncompact. Hence, Theorem 2.2 and Theorem 2.1 apply and we know that any lattice with a nondivergent A-orbit is almost surely grid-DV_N with respect to the Haar measure of the full torus. We shall see in the next section that this result has significant implications towards Minkowki's conjecture. It is worth noting here that a classification of the divergent A-orbits, due to Margulis, is given in [22]. It is proved there that Ax is divergent if and only if there exists $a \in A$, such that the lattice ax is contained in $\mathbb{Q}^d$. There are many extra assumptions one can impose on a lattice x to ensure that x is almost surely grid-DV_N with respect to any nontrivial measure of the form λ_{w+U}. The following two corollaries are examples of such. The first is proved in §5.

Corollary 2.4. *Let $x \in X_d$ be a lattice which does not contain any vector lying on the hyperplanes of the axes (that is the hyperplanes orthogonal to the standard basis vectors). Then for any subspace $\{0\} \neq U < \mathbb{R}^d$, rational with respect to x, and for any $w \in \mathbb{R}^d$, x is λ_{w+U}-almost surely grid-DV_N.*

Corollary 2.5. *Let $x \in X_d$ be a lattice such that there exists a sequence*

$$a_n = \mathrm{diag}\left(e^{t_1^{(n)}}, \ldots, e^{t_d^{(n)}}\right) \in A,$$

such that $a_n x$ converges and for any $1 \leq i \leq d$, $t_i^{(n)}$ diverges. Then x is almost surely grid-DV_N with respect to any nontrivial measure λ_{w+U}

Proof. This follows immediately from Theorem 2.2 as under our assumptions the sequence a_n is not almost finite with respect to any nontrivial subspace. $\qquad\square$

Thanks to the classification of divergent A-orbits we have the following theorem which shows what happens for lattices with divergent orbits.

Theorem 2.6. *Let $x \in X_d$ be a lattice with a divergent A-orbit. Then $V_N(y)$ is discrete (and moreover $\overline{V_N(y)} = V_N(y)$) for any grid $y \in \pi^{-1}(x)$.*

Proof. As the value set does not change along the orbit we conclude from the classification of divergent A-orbits mentioned above, that we may assume that $x \subset \mathbb{Q}^d$. In fact, the reader could verify that the validity of the theorem is stable under commensurability; that is, if x, x' are commensurable lattices (i.e. $x \cap x'$ is of finite index in both), then the sets $V_N(y)$ are discrete for any $y \in \pi^{-1}(x)$ if and only if the same is true for any $y \in \pi^{-1}(x')$. This enables us to assume that $x = \mathbb{Z}^d$. So, we are left to verify that for any vector $v \in \mathbb{R}^d$, the set

$$V_N(v + \mathbb{Z}^d) = \left\{ \prod_1^d (n_i + v_i) : n_i \in \mathbb{Z} \right\},$$

is discrete. To prove this, let $\prod_1^d (n_i^{(k)} + v_i)$ be a converging sequence of elements of $V_N(v + \mathbb{Z}^d)$ and assume by way or contradiction that its elements are distinct (this rules out the possibility of having $n_i^{(k)} + v_i = 0$ for more than one k). For each i, the sequence $(n_i^{(k)} + v_i)$ is discrete. Then we are able to take a subsequence so that for each i, along the subsequence, either $(n_i^{(k)} + v_i)$ is constant and nonzero or it diverges. As we assume the original sequence is converging, we must have that the subsequence is constant from some point which contradicts our assumption. $\qquad\square$

In contrast to the the situation presented in the previous subsection (Theorem 2.3), where nondivergence of the orbit of a lattice under the invariance group was sufficient to ensure that the lattice is almost surely grid-$DV_{P_{n,m}}$ with respect to any measure of the form λ_{w+U}, we work out the following example which shows the existence of a lattice $x \in X_3$, with a nondivergent A-orbit, and a nontrivial subspace $U < \mathbb{R}^3$, rational with respect to x, such that x is not λ_U-almost surely grid-DV_N. In fact, we shall see that the value set $V_N(x + u)$ is discrete for any $u \in U$

Example 2.7. In this example, as we will mix the dimensions, we denote our function N and its invariance group A in dimension d by N_d and A_d respectively. Let $g_0 \in \mathrm{SL}_2(\mathbb{R})$ be such that the lattice $\bar{g}_0$ has a compact orbit in X_2, under the action of the group $A_2 = \{\mathrm{diag}\,(e^t, e^{-t})\}$. It is well known that this is equivalent to saying that the value set $V_{N_2}(\bar{g}_0)$ is discrete and does not contain zero. Let $g = \begin{pmatrix} 1 & 0 \\ 0 & g_0 \end{pmatrix} \in \mathrm{SL}_3(\mathbb{R})$ (here the zeros stand for the corresponding row and column zero vector in $\mathbb{R}^2$), and let $U = \{(*, 0, 0)^t \in \mathbb{R}^3\}$.

Then it is clear that for any grid $y = \bar{g} + u$, $u \in U$ the value set $V_{N_3}(y)$ is discrete. Indeed, if $u = (\alpha, 0, 0)^t$, then

$$V_{N_3}(\bar{g} + u) = \{(n + \alpha)N_2(w) : n \in \mathbb{Z}, w \in \bar{g}_0\},$$

which is clearly discrete as $V_{N_2}(\bar{g}_0)$ is.

To see how this fits with Theorem 2.2 note that if

$$a_n = \operatorname{diag}\left(e^{t_1^{(n)}}, e^{t_2^{(n)}}, e^{-(t_1^{(n)} + t_2^{(n)})}\right)$$

is a sequence in A_3, satisfying that $a_n\bar{g}$ converges in X_3, then it quickly follows that $t_1^{(n)}$ must converge and so the sequence a_n is almost finite with respect to U.

We remark here that for $d \geq 3$, the fact that the invariance group is of higher rank, enabled the author of the present paper to prove [21] the existence of grid-DV$_{N_d}$ lattices. In this case such lattices are also known as GDP lattices (grid-dense-products).

2.4. **Indefinite quadratic forms.** Let $p, q > 0$ be integers such that $d = p + q \geq 3$. Let F be the indefinite quadratic form given by $F(x) = \sum_1^p x_i^2 - \sum_{p+1}^{p+q} x_i^2$. In this case $H_F = \operatorname{SO}(p, q)$. It is not hard to see that F is nondegenerate (in the sense of Definition 1.4) and noncompact. Hence, Theorem 2.2 applies. In fact, as H_F is generated by unipotents, it follows from [16] that H_F has no divergent orbits and so the theorem applies for any lattice.

The fact that H_F is generated by unipotents allows one to obtain much sharper results than Theorem 2.2. In fact, one has the following theorem and corollary which follow from Ratner's orbit closure theorem and Lemma 1.2. We do not elaborate on the arguments as this case is well understood and sharper results than those stated here are available (see for example [17]).

Theorem 2.8. *Let $y \in Y_d$ be given. If the orbit $H_F y$ is closed then the value set $V_F(y)$ is discrete, otherwise the value set is dense.*

It is not hard to see that in this case a grid $y \in Y_d$ has a closed H_F-orbit if and only if $\pi(x) \in X_d$ has a closed orbit and y is *rational*; that is to say $y = x + v$ for $v \in \operatorname{Span}_\mathbb{Q} x$. So we have a full understanding:

Corollary 2.9. *Let $x \in X_d$ be given. If the orbit $H_F x$ is not closed then x is grid-DV$_F$. If the orbit $H_F x$ is closed then rational grids have discrete value sets, while irrational grids are DV$_F$. In any case, x is almost surely grid-DV$_F$ with respect to any nontrivial measure λ_{w+U}.*

3. Applications

We now apply some of the above theorems. In §3.1 we apply Theorem 2.1 to derive new information towards Minkowski's conjecture. In §3.2 we apply Theorem 2.3 to generalize some results on inhomogeneous approximations and provide a partial answer to a problem posed in [23].

3.1. **Minkowski's conjecture.** We shall use the notation of §2.3. The following conjecture is usually attributed to Minkowski and hence named after him. It has remained open for over a century.

Conjecture 3.1 (Minkowski). Let $d \geq 2$ and $N : \mathbb{R}^d \to \mathbb{R}$ be the function $N(x) = \prod_1^d x_i$. Then, for any $y \in Y_d$ one has $V_N(y) \cap [-2^{-d}, 2^{-d}] \neq \emptyset$.

To this date Minkowski's conjecture is verified up to dimension 7. For more information about the interesting history and recent developments we refer the reader to the recent papers [18], [12], where it is proved for dimensions 6 and 7. We say that *a lattice, $x \in X_d$, satisfies Minkowski's conjecture*, if any one of its grids satisfies it. In [3] Bombieri proved that for $d \geq 5$, for any lattice $x \in X_d$ the set $\{y \in \pi^{-1}(x) : y \text{ satisfies Minkowski's conjecture}\}$ has Lebesgue measure $\geq 2^{-\frac{d+1}{2}}$. The results of Narzullaev [20] allows one to strengthen Bombieri's result, still obtaining a poor lower bound. We strengthen this estimate to the best possible from the measure theoretic point of view and prove

Theorem 3.2. *For any $d \geq 2$ and any lattice $x \in X_d$, the set*

$$\left\{y \in \pi^{-1}(x) : y \text{ does not satisfy Minkowski's conjecture}\right\} \tag{3.1}$$

is of Lebesgue measure zero.

Proof. Let $x \in X_d$ be given. We divide the argument into cases. If the orbit Ax of the invariance group is nondivergent, Theorem 2.1 implies that almost any grid y of x is DV_N. In particular, $V_N(y) \cap [-2^{-d}, 2^{-d}] \neq \emptyset$ and the claim follows. If the orbit Ax is divergent, then thanks to the characterization of divergent orbits given in [22] we see that there exists $a \in A$ such that the lattice ax has a basis consisting of vectors with rational coordinates. We refer to such lattices as *rational*. It is well known, (see [15]), that rational lattices satisfy Minkowski's conjecture and therefor, the set in (3.1) is in fact empty. $\square$

We note here that if we impose on the lattice x the further assumptions of Corollaries 2.4, 2.5 then the above theorem can be strengthened to say that the set of possible counterexamples to Minkowski's conjecture in the torus $\pi^{-1}(x)$, has measure zero with respect to any measure of the form λ_{w+U}.

3.2. **Inhomogeneous approximations and nonsingular forms.** In this section we apply Theorem 2.3 to deduce some new results about nonsingular forms. We fix some positive integers n, m and consider a matrix $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$. In Diophantine approximations A is usually regarded as defining a system of n linear forms in m variables, hence we shall refer to such a matrix as a *form*. In first reading it might be useful to take $m = 1$ and n arbitrary, hence $A \in \mathbb{R}^n$ stands for a column vector, or even $n = m = 1$ and then A stands for a real number.

For any dimension k, let $\langle \cdot \rangle$ denote the distance from $\mathbb{Z}^k$ in $\mathbb{R}^k$, i.e. for $v \in \mathbb{R}^k$,

$$\langle v \rangle = \min \left\{ \|v - \mathbf{q}\|_2 : \mathbf{q} \in \mathbb{Z}^k \right\}.$$

Note that for $v \in \mathbb{R}^k$, $\langle v \rangle$ can be thought of as the distance in the torus $\mathbb{T}^k = \mathbb{R}^k / \mathbb{Z}^k$ from v to 0. Here and in the sequel we abuse notation freely and denote points on the k-torus the same as their representatives in $\mathbb{R}^k$.

Definition 3.3. A form $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ is *singular* if for any $\delta > 0$, for any large enough $N \in \mathbb{N}$ one can find $\mathbf{q} \in \mathbb{Z}^m$ such that

$$0 < \|\mathbf{q}\|_2 \le N \ \ and \ \ \langle A\mathbf{q} \rangle < \frac{\delta}{N^{\frac{m}{n}}}. \tag{3.2}$$

By applying Theorem 2.3 we shall deduce some results about nonsingular forms. We shall prove Theorem 3.7 but for the meantime let us formulate a restricted version of it in the form of the following theorem:

Theorem 3.4. *Let $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ be a nonsingular form. Then, for Lebesgue almost any $\boldsymbol{x} \in \mathbb{R}^n$ the following set is dense in $[0, \infty)$,*

$$\left\{ \|\boldsymbol{q}\|_2^m \, \langle A\boldsymbol{q} + \boldsymbol{x} \rangle^n : \boldsymbol{q} \in \mathbb{Z}^m \right\}. \tag{3.3}$$

In particular, the following statement holds

$$\textit{For Lebesgue almost any } \boldsymbol{x} \in \mathbb{R}^n, \ \ \liminf_{\boldsymbol{q} \in \mathbb{Z}^m} \|\boldsymbol{q}\|_2^{\frac{m}{n}} \, \langle A\boldsymbol{q} + \boldsymbol{x} \rangle = 0. \tag{3.4}$$

Remarks 3.5. (1) It follows from the main theorem in [6] that for any nonsingular form $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ the following statement holds

$$\text{For Lebesgue almost any } \mathbf{x} \in \mathbb{R}^n, \ \liminf_{\mathbf{q} \in \mathbb{Z}^m} \|\mathbf{q}\|_2^w \, \langle A\mathbf{q} + \mathbf{x} \rangle = \begin{cases} 0 & \text{for } w < \frac{m}{n} \\ \infty & \text{for } w > \frac{m}{n}. \end{cases} \tag{3.5}$$

The innovation in Theorem 3.4 is that the $\liminf$ in the critical exponent $\frac{m}{n}$ equals zero (and in fact much more, the set in (3.3) is dense).

(2) The above theorem generalizes a series of results. When one takes $n = m = 1$ the number A is nonsingular if and only if it is irrational. In this case Kim [13] proved the validity of (3.4) (generalizing a result of Kurzweil [14]). Tseng reproved Kim's result [23] and raised the following question aiming to generalize Kim's result to higher dimensions:

Question (Tseng). For $m = 1$, hence $A \in \mathbb{R}^n$ is a column vector, is it true that if A does not lie in a rational hyperplane then (3.4) holds.

Galatolo and Peterlongo [11] constructed a (singular) vector in $\mathbb{R}^2$ giving a negative answer to Tseng's question. Nevertheless, Theorem 3.4 tells us that the answer to Tseng's question is positive for nonsingular vectors

(3) Following Baker [1], we say that the form A is *highly singular* if there exists $\epsilon > 0$ such that for all large enough $N \in \mathbb{N}$ one can find $\mathbf{q} \in \mathbb{Z}^m$ such that

$$\|\mathbf{q}\|_2 \le N \ \ and \ \ \langle A\mathbf{q} \rangle < N^{-\frac{m}{n} - \epsilon}. \tag{3.6}$$

It is clear that a highly singular form is singular. We remark here that any vector $A \in \mathbb{R}^n$ which does not lie in a rational hyperplane such that the transposed form A^t is highly singular provides a negative answer to the above quesiton. To see this note that from the main result in [6] it follows that in this case the value of the $\liminf$ in (3.4) equals ∞ for almost any $\mathbf{x}$.

We turn now to the statement and proof of Theorem 3.7. This theorem implies Theorem 3.4 when applied with respect to Lebesgue measure but is in fact considerably stronger. Nevertheless, it is a simple application of Theorem 2.3. In order to state this theorem we need to link the above discussion with the discussion of §2.2, namely the discussion of values of forms on grids. Let $d = n + m$ and recall the notation of §2.2: We write vectors in $\mathbb{R}^d$ as $(\mathbf{x}, \mathbf{y})^t$, where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. We let $P_{n,m} : \mathbb{R}^d \to \mathbb{R}$ be the function given by $P_{n,m}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|_2^n \cdot \|\mathbf{y}\|_2^m$, and denote the invariance group of it by $H_{n,m}$. Recall also that $H_{n,m} = \mathrm{SO}_n(\mathbb{R}) \times \mathrm{SO}_m(\mathbb{R}) \times \{a_{n,m}(t)\}_{t \in \mathbb{R}}$, where $a_{n,m}(t)$ is given by (2.1). For a form $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ we denote

$$g_A = \begin{pmatrix} I_n & A \\ 0 & I_m \end{pmatrix} \in G, \tag{3.7}$$

where I_n, I_m denote the identity matrices of the corresponding dimensions and 0 the zero matrix in $\mathrm{Mat}_{m \times n}(\mathbb{R})$. We denote the lattice $\bar{g}_A \in X_d$ spanned by the columns of g_A by x_A. The following result of Dani is well known (see Theorem 2.14 in [7]) and furnishes the link between the nondivergence condition in Theorem 2.3 and the nonsingularity property.

Lemma 3.6. *A form $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ is singular, if and only if, the orbit $H_{n,m} x_A$ is divergent in X_d.*

Before stating Theorem 3.7 we introduce some more notation. We denote for a subspace $U < \mathbb{R}^d$ and a vector $w_0 \in \mathbb{R}^d$ by $\tilde{\lambda}_{w_0+U}$ the natural U-invariant measure supported on the affine subspace $w_0 + U$. In relation to the notation introduced in §1.2, we note that when U is rational with respect to a lattice x, then the measure λ_{w_0+U} supported on the coset $w_0 + (U + x)$, is obtained from $\tilde{\lambda}_{w_0+U}$ by restricting it to a suitable 'fundamental domain' (which is identified with the coset) and rescaling it to be a probability measure. In particular, a set in the torus $\pi^{-1}(x)$ is λ_{w_0+U}-null if and only if the corresponding set in $\mathbb{R}^d$ is $\tilde{\lambda}_{w_0+U}$-null. Finally, let $p : \mathbb{R}^d \to \mathbb{R}^n$ be the projection $p(\mathbf{x}, \mathbf{y}) = \mathbf{x}$.

Theorem 3.7. *For any nonsingular form $A \in \mathrm{Mat}_{n \times m}(\mathbb{R})$ and for any measure of the form $\mu = p_*(\tilde{\lambda}_{w_0+U})$, where $U \neq \{0\}$ is a subspace of $\mathbb{R}^d$, rational with respect to the lattice x_A, and w_0 is arbitrary, for μ-almost any $\mathbf{x} \in \mathbb{R}^n$*

$$\text{the set} \quad \{\|\mathbf{q}\|_2^m \langle A\mathbf{q} + \mathbf{x} \rangle^n : \mathbf{q} \in \mathbb{Z}^m\} \quad \text{is dense in } [0, \infty). \tag{3.8}$$

Proof. Theorem 2.3 together with Lemma 3.6 imply that x_A is λ_{w_0+U}-a.s. grid-$\mathrm{DV}_{P_{n,m}}$. This means that for $\tilde{\lambda}_{w_0+U}$-almost any $w = (\mathbf{x}, \mathbf{y})^t \in \mathbb{R}^d$ the grid $x_A + w$ is $\mathrm{DV}_{P_{n,m}}$. Calculating the value set we get

$$V_{P_{n,m}}(x_A + w) = \{\|\mathbf{q} + \mathbf{y}\|_2^m \|A\mathbf{q} + \mathbf{x} + \mathbf{p}\|_2^n : \mathbf{p} \in \mathbb{Z}^n, \mathbf{q} \in \mathbb{Z}^m\}. \tag{3.9}$$

Given $w = (\mathbf{x}, \mathbf{y})^t$ such that $x_A + w$ is $\mathrm{DV}_{P_{n,m}}$, for any $\gamma \in [0, \infty)$ there are appropriate sequences $\mathbf{p}_i, \mathbf{q}_i$ satisfying $\lim_i \|\mathbf{q}_i + \mathbf{y}\|_2^m \|A\mathbf{q}_i + \mathbf{x} + \mathbf{p}_i\|_2^n = \gamma$ (and the convergence is not trivial; i.e. there are arbitrarily large i's with $\gamma \neq \|\mathbf{q}_i + \mathbf{y}\|_2^m \|A\mathbf{q}_i + \mathbf{x} + \mathbf{p}_i\|_2^n$). Then, $\|\mathbf{q}_i\|_2$ must go to infinity which in turn implies that $\|A\mathbf{q}_i + \mathbf{x} + \mathbf{p}_i\|_2 \to 0$ and in particular,

$\|A\mathbf{q}_i + \mathbf{x} + \mathbf{p}_i\|_2 = \langle A\mathbf{q}_i + \mathbf{x}\rangle$. It follows that $\lim_i \|\mathbf{q}_i\|_2^m \langle A\mathbf{q}_i + \mathbf{x}\rangle^n = \gamma$ as well. We conclude that

$$p^{-1}\left(\{\mathbf{x} \in \mathbb{R}^n : (3.8) \text{ does not hold}\}\right) \subset$$

$$\left\{w = (\mathbf{x}, \mathbf{y})^t \in \mathbb{R}^d : \text{ the grid } x_A + w \text{ is not } \text{DV}_{P_{n,m}}\right\}.$$

As the right hand side is $\tilde{\lambda}_{w_0+U}$-null the left hand side is $\tilde{\lambda}_{w_0+U}$-null too and by definition, for $p_*(\tilde{\lambda}_{w_0+U})$-almost any $\mathbf{x}$ (3.8) holds. $\qquad\square$

The above theorem, when applied with $U = \mathbb{R}^d$, gives Theorem 3.4 as mentioned earlier. Let us demonstrate the strength of Theorem 3.7 with the following

Corollary 3.8. *Take $m = 1$ and n arbitrary. For any nonsingular vector $v \in \mathbb{R}^n$ and for Lebesgue almost any $t \in \mathbb{R}$ the following set*

$$\{|q| \langle (q+t)v\rangle^n : q \in \mathbb{Z}\},$$

is dense in $[0, \infty)$.

Proof. Apply Theorem 3.7 with $w_0 = 0$ and U being the one dimensional subspace

$$U = \left\{t\begin{pmatrix} v \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1} : t \in \mathbb{R}\right\}.$$

$\qquad\square$

4. The Mixing and the Coset lemmas

In this section we study the following question. Given a sequence of endomorphisms of the d-torus $\gamma_n : \mathbb{T}^d \to \mathbb{T}^d$ and a Haar measure λ_U of a subtorus (see the notation of §1.2), what can one say about the closure $\overline{\{\gamma_n v : n \in \mathbb{N}\}}$ for λ_U-almost any point $v \in \mathbb{T}^d$. We shall see in the Coset lemma below, that unless an obvious obstacle is present this closure must contain a coset of a nontrivial subtorus. In the course of proving the Coset lemma, which is the goal of this section, we shall prove the Mixing lemma, stated below, which is of independent interest on its own. Both the Coset and the Mixing lemmas are stated in somewhat greater generality than we actually need in practice. This is dictated by the argument we use to derive the Coset lemma, which is inductive, hence we need to let the dimensions of the tori to be arbitrary.

Given a compact metric space X, the space of Borel probability measures on it, $\mathcal{P}(X)$, is compact with respect to the weak* topology. When considering measures we shall only refer to the weak* topology. If $\gamma : X \to Y$ is a measurable map between compact metric spaces, it defines a map $\gamma_* : \mathcal{P}(X) \to \mathcal{P}(Y)$, given by $\gamma_*(\eta)(A) = \eta(\gamma^{-1}(A))$ for any Borel measurable $A \subset Y$ and $\eta \in \mathcal{P}(X)$. The following definition is new as far as we know.

Definition 4.1. Given two probability measures $\eta \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ and a sequence of measurable maps $\gamma_n : X \to Y$, we say that γ_n *mixes* η *to* ν if for any probability measure η' which is absolutely continuous with respect to η, one has $(\gamma_n)_*\eta' \to \nu$ weak*. In this case we sometimes say that $(\gamma_n)_*\eta$ converges mixingly to ν.

Note that the above definition is equivalent to the requirement that for any $g \in L^1(\eta)$, and any continuous $f \in C(Y)$ one has

$$\int_X f(\gamma_n(x))g(x)d\eta(x) \to \int_Y f(y)d\nu(y) \int_X g(x)d\eta(x). \tag{4.1}$$

To explain the terminology, recall that given $\nu \in \mathcal{P}(X)$ and a sequence of measurable maps $\gamma_n : X \to X$ preserving ν, the sequence γ_n is said to be a mixing sequence with respect to ν if for any two measurable sets $A, B \subset X$, one has $\nu(\gamma_n(A) \cap B) \to \nu(A)\nu(B)$. This is easily seen to be equivalent to the fact that γ_n mixes ν to itself using the terminology introduced above.

We note here that the fact that γ_n mixes η to ν is much stronger than the fact that $(\gamma_n)_*\eta \to \nu$. In particular, one consequence that we will be interested in is the following simple lemma.

Lemma 4.2. *If $\gamma_n : X \to Y$ mixes η to ν, then for η-almost any $x \in X$ the closure $\overline{\{\gamma_n x : n \in \mathbb{N}\}}$ contains $\mathrm{supp}(\nu)$.*

Proof. Assume that the conclusion is false. Then, there must exist a bump function $f \in C(Y)$ (i.e. $0 \le f \le 1$) with $\int_Y f d\nu > 0$ and a measurable set $\Omega \subset X$ with $\eta(\Omega) > 0$ such that for any $x \in \Omega$, the sequence $\gamma_n x$ never visits $U = \{y \in Y : f(y) > 0\}$. Let $g \in L^1(\eta)$ be the characteristic function of Ω. We now see that the left hand side of (4.1) is constantly zero while the right hand side is positive which is the desired contradiction. $\square$

Lemma 4.3 (Mixing lemma). *Let $U < \mathbb{R}^{d_1}$ be a subspace rational with respect to $\mathbb{Z}^{d_1}$. Let $\gamma_n : \mathbb{T}^{d_1} \to \mathbb{T}^{d_2}$ be a sequence of homomorphisms induced by the matrices $\gamma_n \in \mathrm{Mat}_{d_1 \times d_2}(\mathbb{Z})$. Assume the sequence $(\gamma_n)_*\lambda_U$ converges weak* to $\mu \in \mathcal{P}(\mathbb{T}^{d_2})$. Then, there exists a subspace $V < \mathbb{R}^{d_2}$, rational with respect to $\mathbb{Z}^{d_2}$, such that $\mu = \lambda_V$. Moreover, the sequence γ_n mixes λ_U to λ_V if and only if there are no integer vectors $\mathbf{m} \in \mathbb{Z}^{d_2}$ and $0 \ne \mathbf{k} \in U$ satisfying $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ for infinitely many n's.*

Before turning to the proof we make some clarifying remarks and work out some examples. First, the fact that μ is a Haar measure is not the essence of the lemma. It is an exercise to prove that any weak* limit of Haar measures of subtori is again a Haar measure of a subtorus (this fact is implicitly used in the proof below). The true content of the lemma is the fact that in the absence of the obvious obstacle to mixing – the existence of $\mathbf{m}, \mathbf{k} \ne 0$ such that $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ infinitely often – γ_n actually mixes λ_U to λ_V, which is significantly stronger than the convergence of $(\gamma_n)_*\lambda_U$ to λ_V. Second, note that given any sequence $\gamma_n : \mathbb{T}^{d_1} \to \mathbb{T}^{d_2}$ of homomorphisms and a Haar measure λ_U of a subtorus $U + \mathbb{Z}^{d_1} \subset \mathbb{T}^{d_1}$, one can always assume, after passing to a subsequence, that $(\gamma_n)_*\lambda_U$ converges by the compactness of $\mathcal{P}(\mathbb{T}^{d_2})$. Third, note that one could rephrase the condition ensuring mixing as saying that for any $\mathbf{m} \in \mathbb{Z}^{d_2}$, either $\gamma_n^t \mathbf{m}$ is eventually in $U^\perp$ or the distance from $\gamma_n^t \mathbf{m}$ to $U^\perp$ goes to ∞. Finally we note that when $d_1 = d_2 = 2$, $U = \mathbb{R}^{d_1}$, and the γ_n's are automorphisms, the lemma is proved in [2] Lemma 2.2.

Example 4.4. Let $\gamma \in \mathrm{GL}_d(\mathbb{Z})$ ($d \ge 2$) be an automorphism of the d-torus which has an irreducible characteristic polynomial all of whose roots are real. Then for any subtorus

$U + \mathbb{Z}^d$, the sequence $\gamma_*^n \lambda_U$ converges mixingly to the Haar measure of the full d-torus. In particular, for λ_U almost any point $v \in \mathbb{T}^d$ the orbit $\{\gamma^n v\}$ is dense in $\mathbb{T}^d$.

We split the argument into two parts. In the first we show that if a subsequence $\gamma_*^{n_i} \lambda_U$ converges to some λ_V, then $V = \mathbb{R}^d$, which proves that $\gamma_*^n \lambda_U$ converges to $\lambda_{\mathbb{R}^d}$. In the second we show that the convergence is mixing. We rely on the following two properties of γ (resp. γ^t):

(1) γ (resp. γ^t) is diagonalizable over $\mathbb{R}$ with all eigenvalues distinct and of absolute value $\neq 1$. Hence, for any $v \in \mathbb{R}^d$, $\gamma^n v$ converges projectively to a one dimensional eigenspace of γ.

(2) Any γ-invariant (resp. γ^t-invariant) subspace (which is necessarily a direct sum of eigenspaces) is not a subspace of any proper rational subspace (this follows from the irreducibility of the characteristic polynomial).

First step: If n_i is such that $\gamma_*^{n_i} \lambda_U$ converges to some λ_V, then in particular, for any $v \in U$, any projective limit of $\gamma^{n_i} v$ is a line in V. But from (1) it follows that such a line is an eigenspace of γ and property (2) implies that V, which is a rational subspace, must equal $\mathbb{R}^d$.

Second step: We need to show that there could not exist integer vectors $\mathbf{m} \in \mathbb{Z}^d, 0 \neq \mathbf{k} \in U$ such that $(\gamma^t)^n \mathbf{m} + \mathbf{k} \in U^\perp$ for infinitely many n's. From property (1) we deduce that $(\gamma^t)^n \mathbf{m}$ converges projectively to a 1-dimensional eigenspace of γ^t which is contained in $U^\perp$ (as $\mathbf{k}$ is fixed and $(\gamma^t)^n \mathbf{m}$ must diverge) which is a rational proper subspace (as $U \neq \{0\}$). This contradicts property (2).

Example 4.5. Let $\gamma_n = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$. If $U = \{(x, x) \in \mathbb{R}^2 : x \in \mathbb{R}\}$, then although $(\gamma_n)_* \lambda_U$ converges to the Haar measure of the full 2-torus, γ_n does not mix λ_U to $\lambda_{\mathbb{R}^2}$. This is because, if we choose $\mathbf{m} = (1, 0)^t$ then $\mathbf{m}$ is a fixed point of γ_n^t and is not in $U^\perp$, so we see that $\gamma_n^t \mathbf{m}$ is not eventually in $U^\perp$ and its distance from $U^\perp$ does not diverge. Indeed, Lemma 4.2 does not apply and for λ_U-almost any $v \in \mathbb{T}^2$ the closure $\overline{\{\gamma_n v\}}$ equals a coset of a lower dimensional subtorus.

Proof of the Mixing lemma. For any d and $\mathbf{m} \in \mathbb{Z}^d$, let $e_{\mathbf{m}}(v) = e^{2\pi i <\mathbf{m}, v>} : \mathbb{T}^d \to \mathbb{C}$. Note that in our notation there is no reference to the dimension and the reader should understand from the context what is the domain of the character $e_{\mathbf{m}}$. In particular, there will be times in which in one equation two dimensions will be mixed. We need to find a rational subspace $V < \mathbb{R}^{d_2}$ such that (4.1) holds for any $f \in C(\mathbb{T}^{d_2}), g \in L^1(\lambda_U)$. A standard argument shows that it is enough to verify the validity of (4.1) when f and g are chosen from sets spanning dense subspaces of $C(\mathbb{T}^{d_2})$ and $L^1(\lambda_U)$ respectively. From the Stone-Weierstrass theorem it follows that $\{e_{\mathbf{m}} : \mathbf{m} \in \mathbb{Z}^{d_2}\}$ spans a dense subspace of $C(\mathbb{T}^{d_2})$. Also, the set $\{e_{\mathbf{k}} : \mathbf{k} \in U\}$ spans a dense subspace in $L^1(\lambda_U)$. This is because it forms an orthonormal basis to $L^2(\lambda_U)$, which is dense in $L^1(\lambda_U)$ by the Cauchy Schwarz inequality.

14 URI SHAPIRA

We conclude that we need to find a rational subspace $V < \mathbb{R}^{d_2}$ such that the following convergence holds

For any $\mathbf{m} \in \mathbb{Z}^{d_2}, \mathbf{k} \in U, \quad \int_{\mathbb{T}^{d_1}} e_{\gamma_n^t \mathbf{m} + \mathbf{k}}(v) d\lambda_U(v)$ \hfill (4.2)

$$= \int_{\mathbb{T}^{d_1}} e^{2\pi i <\gamma_n^t \mathbf{m} + \mathbf{k}, v>} d\lambda_U(v) = \int_{\mathbb{T}^{d_1}} e_{\mathbf{m}}(\gamma_n v) e_k(v) d\lambda_U(v) \quad (4.3)$$

$$\to \int_{\mathbb{T}^{d_2}} e_{\mathbf{m}}(v) d\lambda_V(v) \int_{\mathbb{T}^{d_1}} e_{\mathbf{k}}(v) d\lambda_U(v). \quad (4.4)$$

Given a rational subspace $V < \mathbb{R}^d$ and a corresponding Haar measure λ_V, a short calculation shows that for any $\mathbf{m} \in \mathbb{Z}^d$ we have

$$\int e_{\mathbf{m}}(v) d\lambda_V(v) = \left\{ \begin{array}{ll} 1 & \text{if } \mathbf{m} \in V^\perp, \\ 0 & \text{otherwise.} \end{array} \right. \quad (4.5)$$

Working with equation (4.5), we conclude that for any choice of V the values in (4.2), (4.4) satisfy

$$(4.2) = \left\{ \begin{array}{ll} 1 & \text{if } \gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp, \\ 0 & \text{otherwise} \end{array} \right. ; \quad (4.4) = \left\{ \begin{array}{ll} 1 & \text{if } \mathbf{m} \in V^\perp \text{ and } \mathbf{k} = 0, \\ 0 & \text{otherwise} \end{array} \right.$$

We conclude that given λ_V, the sequence γ_n mixes λ_U to λ_V if and only if the values of (4.2) and (4.4) agree for all large enough n's. This implies immediately the only if part of the lemma; if there exist $\mathbf{m} \in \mathbb{Z}^{d_2}, 0 \neq \mathbf{k} \in U$ such that $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ for infinitely many n's, then for these values of n the value of (4.2) is 1 while the value of (4.4) is 0.

We are left to prove the if part. Assume then that for any $\mathbf{m} \in \mathbb{Z}^{d_2}$ and $0 \neq \mathbf{k} \in U$, $\gamma_n^t \mathbf{m} + \mathbf{k}$ is eventually outside $U^\perp$. We shall conclude the proof by showing that for any subsequence of γ_n there exists yet another subsequence $\tilde{\gamma}_n$ and a subspace V, such that $\tilde{\gamma}_n$ mixes λ_U to λ_V. To see why this concludes the proof note first that this implies that $\mu = \lambda_V$, as we assume that $(\gamma_n)_* \lambda_U$ converges to μ, hence any subsequence of it converges to μ too. In particular, this shows that V does not depend on the initial subsequence. Second, if γ_n does not mix λ_U to λ_V, then the above discussion shows that there must exist integer vectors $\mathbf{m} \in \mathbb{Z}^d$ and $\mathbf{k} \in U$ such that for infinitely many n's the value in (4.2) is different from the value in (4.4); we obtain a contradiction once we take as the initial subsequence those γ_n's for which the values in (4.2) and (4.4) do not agree, seeing that there could not be a subsequence $\tilde{\gamma}_n$ of that subsequence which mixes λ_U to λ_V.

To this end, let a subsequence of γ_n be given. By a standard diagonal argument it has a subsequence $\tilde{\gamma}_n$ such that for any $\mathbf{m} \in \mathbb{Z}^{d_2}$, either $\tilde{\gamma}_n^t \mathbf{m} \in U^\perp$ for all but finitely many n's, or $\tilde{\gamma}_n^t \mathbf{m} \notin U^\perp$ for all but finitely many n's. Define

$$V^\perp = \text{span} \left\{ \mathbf{m} \in \mathbb{Z}^{d_2} : \tilde{\gamma}_n^t \mathbf{m} \in U^\perp \text{ for all large } n's \right\} \quad (4.6)$$

$$= \text{span} \left\{ \mathbf{m} \in \mathbb{Z}^{d_2} : \tilde{\gamma}_n^t \mathbf{m} \in U^\perp \text{ for infinitely many } n's \right\}.$$

We remark that the above set of integer vectors spanning $V^\perp$ is clearly a group and in fact it is the intersection of $V^\perp$ with $\mathbb{Z}^{d_2}$; that is, if $\mathbf{m} \in V^\perp$ is an integer vector, then $\tilde{\gamma}_n^t \mathbf{m} \in U^\perp$ for all large enough n's.

We now check that along the subsequence $\tilde{\gamma}_n$ the values of (4.2), (4.4) agree for all large enough n's, hence concluding that $\tilde{\gamma}_n$ mixes λ_U to λ_V as desired. There are two cases in which the values in (4.2), (4.4) do not agree (with γ_n replaced by $\tilde{\gamma}_n$).

(1) Either $\tilde{\gamma}_n^t \mathbf{m} + \mathbf{k} \notin U^\perp$ and $\mathbf{k} = 0$ and $\mathbf{m} \in V^\perp$,
(2) or $\tilde{\gamma}_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ but either $\mathbf{k} \neq 0$ or $\mathbf{m} \notin V^\perp$.

It follows from the remark made after the definition of $V^\perp$, that (1) cannot happen infinitely many times. Assume (2) holds for infinitely many n's. It cannot be that $\mathbf{k} \neq 0$ because of our assumption that $\gamma_n^t \mathbf{m} + \mathbf{k}$ is eventually outside $U^\perp$ for any $\mathbf{m}$ and nonzero $\mathbf{k}$. On the other hand, it cannot be that $\mathbf{k} = 0$ by our construction of $\tilde{\gamma}_n$ and the definition of $V^\perp$, as it would mean that eventually $\tilde{\gamma}_n^t \mathbf{m} \in U^\perp$, hence $\mathbf{m} \in V^\perp$. In any case we arrive to a contradiction as desired. $\square$

We now use the Mixing lemma to prove another lemma which serves as the main tool used to prove Theorem 2.2. As in the case of the mixing lemma this lemma is of independent interest on its own.

Lemma 4.6 (Coset lemma). *Let $U < \mathbb{R}^{d_1}$ be a subspace, rational with respect to $\mathbb{Z}^{d_1}$. Let $\gamma_n : \mathbb{T}^{d_1} \to \mathbb{T}^{d_2}$ be a sequence of homomorphisms induced by the matrices $\gamma_n \in \mathrm{Mat}_{d_1 \times d_2}(\mathbb{Z})$ such that the restrictions $\gamma_n|_U$ form an infinite set. Then there exists a subspace $\{0\} \neq V < \mathbb{R}^{d_2}$, rational with respect to $\mathbb{Z}^{d_2}$, such that for λ_U-almost any $v \in \mathbb{T}^{d_1}$ the closure $\overline{\{\gamma_n v : i \in \mathbb{N}\}}$ contains a coset of the subtorus $V + \mathbb{Z}^{d_2}$*

Note that in the above lemma the subtorus $V + \mathbb{Z}^{d_2}$ is fixed while its coset depend on the initial point.

Proof. Along the argument we will take subsequences a large number of times. We will abuse notation and continue to denote the subsequences by the same symbols. The proof goes by induction on d_2 but we first make some general observations. First, by taking a subsequence we may assume that the restrictions $\gamma_n|_U$ are distinct and that $(\gamma_n)_* \lambda_U$ converges. If there are no integer vectors $\mathbf{m} \in \mathbb{Z}^{d_2}, 0 \neq \mathbf{k} \in U$ such that $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ for infinitely many n's, then the Mixing lemma implies that γ_n mixes λ_U to some λ_V. Of course $V \neq \{0\}$ as equality would imply that $U = \{0\}$ which contradicts our assumption that the restrictions $\gamma_n|_U$ are distinct. Lemma 4.2 now tells us that for λ_U-almost any $v \in \mathbb{T}^{d_1}$, the closure $\overline{\{\gamma_n v\}}$ actually contains the subtorus $V + \mathbb{Z}^{d_2}$ and the lemma follows with the additional information that the coset we gain in the closure does not depend on the initial point. Assume now that there exits integer vectors $\mathbf{m} \in \mathbb{Z}^{d_2}, 0 \neq \mathbf{k} \in U$ such that $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$ for infinitely many n's. By passing to a subsequence we may assume this happens for all n. Note that the validity of the lemma is stable under a change of variable, that is if $A \in \mathrm{SL}_{d_1}(\mathbb{Z}), B \in \mathrm{SL}_{d_2}(\mathbb{Z})$ are two automorphisms of $\mathbb{T}^{d_i}$, then the lemma is true for U, γ_n if and only if it is true for $A^{-1}U, B\gamma_n A$. It follows that we might assume that U is the standard embedding of $\mathbb{R}^r$ in $\mathbb{R}^{d_1}$ for some $1 \leq r \leq d_1$ (i.e. U is the subspace corresponding to the first r coordinates) and that $\mathbf{k} \in U$ is collinear to e_1, i.e. there exists an integer ℓ such that $\mathbf{k} = -\ell e_1 = (-\ell, 0, \ldots, 0)^t \in \mathbb{Z}^{d_1}$. This is done by choosing A properly. By choosing B properly we see that we may assume that $\mathbf{m} = e_1$.

Indeed, $\mathbf{m}$ was chosen to be an integer vector satisfying $\gamma_n^t \mathbf{m} + \mathbf{k} \in U^\perp$. Any choice of B results in replacing the original $\mathbf{m}$ by $(B^t)^{-1}\mathbf{m}$. Choosing B properly, we can guarantee that $\mathbf{m}$ is an integer multiple of e_1, but as it can be chosen to be primitive, we may assume it actually equals e_1. To summarize, if we assume that γ_n does not mix λ_U to some λ_V, we may assume we are in the following position:

(1) The space U equals the standard copy of $\mathbb{R}^r$ in $\mathbb{R}^{d_1}$.
(2) The matrices γ_n satisfy $\gamma_n^t e_1 - \ell e_1 \in U^\perp$. This means that the first row of the matrix γ_n is of the form $(\underbrace{\ell, 0, \ldots, 0}_{r}, * \cdots *)$.

The particular shape of the first row of the γ_n's is what enables us to have a reduction to a lower dimension and invoke the inductive hypothesis. We now describe the proof by induction.

The case $d_2 = 1$: We prove this case by induction on d_1. Note that in this case there is just one possibility for V and only one coset; this means that we need to prove that for λ_U-almost any $v \in \mathbb{T}^{d_1}$, the sequence $\gamma_n v$ is dense in $\mathbb{T}^1$. For $d_1 = 1$, γ_n are simply integer numbers and $U = \mathbb{R}$. The fact that we assume that γ_n are all distinct imply that γ_n always mixes λ_U to itself and we are done as explained above. Assume we know the lemma for $d < d_1$ and $d_2 = 1$. If γ_n mixes λ_U then we are done as explained above. If not, then we may assume after a change of variable, as explained above, that the first coordinate of the matrix γ_n (which is a row vector) is fixed and equals to some integer ℓ. In particular, this forces the dimension of U to be ≥ 2, as we assume that the restrictions $\gamma_n|_U$ are all distinct. The lemma now clearly follows from the inductive hypothesis as given a vector $v = (v_1, \ldots, v_r, 0 \ldots 0)^t \in U$, we know that for almost any $v' = (v_2, \ldots, v_r, 0 \ldots 0)^t$ the sequence $\gamma_n' v'$ is dense in $\mathbb{T}^1$ (here γ_n' are the γ_n's after we erase the first coordinate). But then we see that for λ_U-almost any $v \in U$ the sequence $\gamma_n v = \ell v_1 + \gamma_n' v'$ is dense in $\mathbb{T}^1$. This finishes the case $d_2 = 1$ which is the base of our induction.

The inductive step: Assume we know the lemma when the dimension of the image torus is less than d_2. If γ_n mixes λ_U then we are done as explained above. If not, then after a change of variable we may assume that the γ_n's are of the form

$$\gamma_n = \begin{pmatrix} \ell & s_n \\ p_n & q_n \end{pmatrix}, \tag{4.7}$$

where p_n is a column vector, $s_n = (\underbrace{0, \ldots, 0}_{r-1}, * \cdots *)$, and q_n are matrices of the appropriate dimension. Let us denote by γ_n' the matrices obtained from γ_n by erasing the first row. We view them as homomorphisms from $\mathbb{T}^{d_1}$ to $\mathbb{T}^{d_2-1}$. The restrictions of γ_n' to U (which is the copy of $\mathbb{R}^r$ in $\mathbb{R}^{d_1}$) must all be distinct as we know this for the γ_n's and their first row is identical as far as U is concerned. From the inductive hypothesis we conclude that there exists a nontrivial subtorus $V + \mathbb{Z}^{d_2-1}$ in $\mathbb{T}^{d_2-1}$ such that for λ_U-almost any $v \in \mathbb{T}^{d_1}$ the closure $\overline{\{\gamma_n' v\}}$ contains a coset $w_v' + (V + \mathbb{Z}^{d_2-1})$, where $w_v' \in \mathbb{R}^{d_2-1}$. We conclude that for such v's the closure $\overline{\{\gamma_n v\}}$ contains the coset $w_v + (V + \mathbb{Z}^{d_2})$, where w_v is obtained from w_v' by adding to it as a first coordinate ℓv_1. This concludes the proof. $\square$

We work out a few examples to develop a feeling about the null set of points which corresponds to closures which do not contain a nontrivial coset. A first example is given in Example 4.5, where the points $v = (x, x)^t \in U$ for which the closure of the sequence $\gamma_n v = (x, (n+1)x)$ in $\mathbb{T}^2$ does not contain a coset of the subtorus corresponding to the second coordinate are exactly those with $x \in \mathbb{Q}$ which clearly form a λ_U-null set. As a slightly more complicated example we have

Example 4.7. Let Let $\gamma_n = \begin{pmatrix} 1 & 0 \\ n & n \end{pmatrix}$, and $U = \mathbb{R}^2$. Then for $v = (x, y)^t$ the closure of the sequence $\gamma_n v = (x, n(x+y))$ equals a coset of the subtorus corresponding to the second coordinate if and only if v does not lie on a line of the form $\{v : x + y = q\}$ for some $q \in \mathbb{Q}$. This countable collection of lines is clearly a λ_U-null set but still a dense set of dimension 1.

Example 4.8. Let M be a diagonalizable epimorphism of the d-torus. When we apply the Coset lemma to the case where $\gamma_n = M^n$, it follows that for Lebesgue almost any v the sequence $M^n v$ contains in its closure a full coset of some nontrivial subtorus. Recently it was proved in [5] (see also [4]) that for any given w, the set $\{v \in \mathbb{T}^d : \inf_n \mathrm{d}(M^n v, w) > 0\}$ is dense and of full dimension. In fact, it is shown that this set is a winning set for Schmidt's game (this generalizes Dani's work [8]). In particular, the set of points $v \in \mathbb{T}^d$ which violate the conclusion of the Coset lemma, although is null, is still large from other perspectives.

5. Proofs of Theorem 2.2 and Corollary 2.4

We begin by rephrasing the Coset lemma when we let the lattice vary. We shall use the following notation: For a grid $y = x + w \in Y_d$ and $v \in \mathbb{R}^d$ we denote by $y + v$ the grid $x + (w + v)$. Given a subspace $V < \mathbb{R}^d$ we let $y + V = \{y + v : v \in V\} \subset \pi^{-1}(x)$. Thus $y + V$ is simply the coset of V passing through the grid y in the torus $\pi^{-1}(x)$.

Lemma 5.1. *Let $x_1, x_2 \in X_d$ be two lattices and $h_n \in G$ be such that $h_n x_1 \to x_2$. Let $U < \mathbb{R}^d$ be a rational subspace with respect to x_1 and $w \in \mathbb{R}^d$ be given. Assume that h_n is not almost finite with respect to U (in the sense of Definition 1.5). Then, there exists a subspace $\{0\} \neq V < \mathbb{R}^d$, rational with respect to x_2, such that for λ_{w+U}-almost any grid $y \in \pi^{-1}(x_1)$ the closure $\overline{\{h_n y : n \in \mathbb{N}\}}$ contains a coset $z_y + V$ for some $z_y \in \pi^{-1}(x_2)$.*

Proof. We first note that after passing to a subsequence we may assume that $h_n(w + x_1)$ converges to some grid in $\pi^{-1}(x_2)$. This clearly reduces the lemma to the case where we are dealing with a Haar measure of a subtorus λ_U rather than a translate of it λ_{w+U}.

Let $g_i \in G$ be such that $x_i = g_i \Gamma$ for $i = 1, 2$. There exists a sequence $\epsilon_n \in G$ with $\epsilon_n \to e$ such that $\epsilon_n h_n x_1 = x_2$, or in other words, $\gamma_n = g_2^{-1} \epsilon_n h_n g_1 \in \Gamma$. Let $U' = g_1^{-1} U$. Then U' is rational with respect to $\mathbb{Z}^d$. The fact that h_n is not almost finite with respect to U translates to the fact that the restrictions $\gamma_n|_{U'}$ form an infinite set. Hence, the Coset lemma applies and gives us the existence of a subspace $\{0\} \neq V' < \mathbb{R}^d$, rational with respect to $\mathbb{Z}^d$, such that for $\lambda_{U'}$-almost any grid y' of $\mathbb{Z}^d$ the closure of the sequence

$\gamma_n y'$ contains a coset of the subtorus $V' + \mathbb{Z}^d$. This translates to the fact that for $V = g_2 V'$ and λ_U-almost any grid $y \in \pi^{-1}(x_1)$, the closure of the sequence $\epsilon_n h_n y$, contains a coset of the subtorus $V + x_2$. This implies the same statement for the sequence $h_n y$, as $\epsilon_n \to e$, and the lemma follows. $\qquad\square$

Before turning to the proof of Theorem 2.2 we note that in the terminology introduced at the beginning of this section, we have that a continuous function $F : \mathbb{R}^d \to \mathbb{R}$ is nondegenerate in the sense of Definition 1.4 if and only if for any grid $y \in Y_d$ and any subspace $\{0\} \neq V < \mathbb{R}^d$ one has $\cup_{y' \in \{y+V\}} V_F(y') = F(\mathbb{R}^d)$.

Proof of Theorem 2.2. Let F, x, U, w, and h_n be as in the statement. Denote $x = x_1$ and $\lim h_n x_1 = x_2$. Lemma 5.1 implies that there exists a nontrivial subspace $V < \mathbb{R}^d$ such that for λ_{w+U}-almost any $y \in \pi^{-1}(x_1)$ the closure $\overline{H_F y}$ contains a full coset $z_y + V$, for $z_y \in \pi^{-1}(x_2)$. The Inheritance lemma (Lemma 1.2) implies now that such a grid satisfies

$$V_F(y) \supset \cup_{y' \in \{z_y + V\}} V_F(y') = F(\mathbb{R}^d),$$

which concludes the proof. $\qquad\square$

Proof of Corollary 2.4. Let x, U, and w be as in the statement. From the classification of divergent H_F-orbits (see [22] and §2) we see that as x does not contain any vectors on the axes, it has a nondivergent H_F-orbit. Let $h_n \in H_F$ be a diverging sequence such that $h_n x$ converges. We only need to argue why the sequence h_n cannot be almost finite with respect to U. Indeed, assume that h_n is almost finite with respect to U. After passing to a subsequence and possibly permuting the coordinates (which we ignore in order to ease our notation) we can assume that $h_n = \text{diag}\left(e^{t_1^{(n)}}, \ldots, e^{t_d^{(n)}}\right)$, where for some $1 \leq r \leq d$ we have that $t_i^{(n)}$ diverge for $i \leq r$ and converge for $i > r$. It follows that as h_n is almost finite with respect to U, we must have $U < \left\{(\underbrace{0, \ldots, 0}_{r}, * \cdots *)^t \in \mathbb{R}^d\right\}$. This leads to a contradiction as U is rational with respect to x and we assumed that x does not contain any point on the hyperplanes of the axes. $\qquad\square$

References

[1] R. C. Baker. Singular n-tuples and Hausdorff dimension. *Math. Proc. Cambridge Philos. Soc.*, 81(3):377–385, 1977.

[2] V. Bergelson and A. Gorodnik. Weakly mixing group actions: a brief survey and an example. In *Modern dynamical systems and applications*, pages 3–25. Cambridge Univ. Press, Cambridge, 2004.

[3] Enrico Bombieri. Alcune osservazioni sul prodotto di n forme lineari reali non omogenee. *Ann. Math. Pura Appl. (4)*, 61:279–285, 1963.

[4] R. Broderick, Y. Bugeaud, L. Fishman, D. Kleinbock, and B. Weiss. Schmidt's game, fractals, and numbers normal to no base. Preprint, available on arXiv.

[5] R. Broderick, L. Fishman, and D. Kleinbock. Schmidt's game, fractals, and orbits of toral endomorphisms. Preprint, available on arXiv.

[6] Yann Bugeaud and Michel Laurent. On exponents of homogeneous and inhomogeneous Diophantine approximation. *Mosc. Math. J.*, 5(4):747–766, 972, 2005.

[7] S. G. Dani. Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.*, 359:55–89, 1985.

[8] S. G. Dani. On orbits of endomorphisms of tori and the Schmidt game. *Ergodic Theory Dynam. Systems*, 8(4):523–529, 1988.

[9] H. Davenport. Indefinite binary quadratic forms, and Euclid's algorithm in real quadratic fields. *Proc. London Math. Soc. (2)*, 53:65–82, 1951.

[10] M. Einsiedler and J Tseng. Badly approximable systems of affine forms, fractals, and Schmidt games. Preprint.

[11] S. Galatolo and P. Peterlongo. Long hitting time, slow decay of correlations and arithmetical properties. Preprint, available on arXiv.

[12] R. J. Hans-Gill, Madhu Raka, and Ranjeet Sehmi. On conjectures of Minkowski and Woods for $n = 7$. *J. Number Theory*, 129(5):1011–1033, 2009.

[13] Dong Han Kim. The shrinking target property of irrational rotations. *Nonlinearity*, 20(7):1637–1643, 2007.

[14] J. Kurzweil. On the metric theory of inhomogeneous diophantine approximations. *Studia Math.*, 15:84–112, 1955.

[15] A. M. Macbeath. Factorization of matrices and Minkowski's conjecture. *Proc. Glasgow Math. Assoc.*, 5:86–89 (1961), 1961.

[16] G. A. Margulis. On the action of unipotent groups in the space of lattices. In *Lie groups and their representations (Proc. Summer School, Bolyai, János Math. Soc., Budapest, 1971)*, pages 365–370. Halsted, New York, 1975.

[17] G. A. Margulis and A. Mohammadi. Quantitative version of the Oppenheim conjecture for inhomogeneous quadratic forms. Preprint, available on arXiv.

[18] Curtis T. McMullen. Minkowski's conjecture, well-rounded lattices and topological dimension. *J. Amer. Math. Soc.*, 18(3):711–734 (electronic), 2005.

[19] Calvin C. Moore. Ergodicity of flows on homogeneous spaces. *Amer. J. Math.*, 88:154–178, 1966.

[20] H. N. Narzullaev. The product of linear inhomogeneous forms. *Mat. Zametki*, 16:365–374, 1974.

[21] Uri Shapira. A solution to a problem of Cassels and diophantine properties of cubic numbers. To appear in Ann. of Math., available on arXiv at http://arxiv.org/abs/0810.4289v2.

[22] George Tomanov and Barak Weiss. Closed orbits for actions of maximal tori on homogeneous spaces. *Duke Math. J.*, 119(2):367–392, 2003.

[23] Jimmy Tseng. On circle rotations and the shrinking target properties. *Discrete Contin. Dyn. Syst.*, 20(4):1111–1122, 2008.

[24] Robert J. Zimmer. *Ergodic theory and semisimple groups*, volume 81 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1984.